DOCUMENT RESUME

ED 243 964 TM 840 287

AUTHOR Bliss, Leonard B.
TITLE Item Analysis and Report of Student Skills of Secondary School Students. Technical Report #2, Public School Basic Skills Achievement Survey.
INSTITUTION College of the Virgin Islands, St. Thomas. Caribbean Research Inst.
PUB DATE Jun 82
NOTE 103p.; For Technical Report No. 1, see ED 227 150.
AVAILABLE FROM Caribbean Research Institute, College of the Virgin Islands, St. Thomas, Virgin Islands 00801 ($4.00).
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC05 Plus Postage.
DESCRIPTORS *Academic Achievement; *Achievement Tests; *Basic Skills; Difficulty Level; Elementary Secondary Education; *Item Analysis; Language Skills; Mathematics Skills; Reading Skills; *School Districts; Test Items
IDENTIFIERS Stanford Achievement Tests; Stanford Test of Academic Skills; *Virgin Islands

ABSTRACT

A sample of slightly over 1,500 students from even-numbered grades in public schools of the U.S. Virgin Islands were given the 1973 edition of the Stanford Achievement Test (in grades 2, 4, 6, and 8) and the Test of Academic Skills (grades 10 and 12) in an attempt to assess student academic achievement in the basic skill areas of mathematics, reading, and English Language. This report describes the data analysis which involved a detailed item analysis of each item on each test given to the sample of students in grades 8, 10, and 12 as well as a summary of student skills based on their achievement along objectives provided by the test publisher and keyed to individual test items. Measures of item difficulty and item discrimination were calculated for the entire territorial school system and for the individual school districts. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED243964

Virgin Islands of the United States Public School Basic Skills Achievement Survey

TECHNICAL REPORT #2 ITEM ANALYSIS AND REPORT OF STUDENT SKILLS OF SECONDARY SCHOOL STUDENTS

by

Leonard B. Bliss, Ph.D.

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

L. B. Bliss

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 840 287

CARIBBEAN RESEARCH INSTITUTE
CHARLOTTE AMALIE, ST. THOMAS
U.S. VIRGIN ISLANDS 00802

Virgin Islands of the United States
Public School Basic Skills
Achievement Survey

Technical Report # 2
Item Analysis and Report of Student Skills of Secondary School Students

Caribbean Research Institute, College of the Virgin Islands
Leonard B. Bliss, Ph.D. - Principal Investigator
June 1982

Contents

	Page
Abstract	2
Introduction	3
Background	5
Item analysis	6
Difficulty indices	6
Discrimination indices	7
Summary of student skills	11
Availability of data	13
8th Grade (Advanced Level)	14
Vocabulary	16
Reading Comprehension	18
Mathematics Concepts	29
Mathematics Computations	32
Mathematics Applications	36
Spelling	40
Language	43
10th Grade (TASK Level I)	51
Reading	53
English	63
Mathematics	67
12th Grade (TASK Level II)	73
Reading	75
English	84
Mathematics	88
Discussion	96
References	100

Abstract

A sample of slightly over 1500 students from even numbered grades in public schools of the U.S. Virgin Islands were given the 1973 edition of the Stanford Achievement Test (in grades 2, 4, 6, and 8) and the Test of Academic Skills (grades 10 and 12) in an attempt to assess student academic achievement in the basic skill areas of mathematics, reading, and English Language. This report describes Phase II of the data analysis which involved a detailed item analysis of each item on each test given to the sample of students in grades 8, 10, and 12 as well as a summary of student skills based on their achievement along objectives provided by the test publisher and keyed to individual test items. Measures of item difficulty and item discrimination were calculated for the entire territorial school system and for the individual school districts.

At the request of the U•S• Virgin Islands Department of Education and the Board of Trustees of the College of the Virgin Islands, the Caribbean Research Institute embarked on a study of basic skills achievement in U•S• Virgin Islands public schools. It soon became clear that any strategy designed to improve basic skills achievement needed to start off with a fairly detailed description of current achievement levels of students in territorial public schools. It was found that this was not available.

The task force set up to design the study decided that the most efficient way to obtain information on levels of basic skills achievement was to administer a standardized achievement test to a representative sample of students and to analyze the results of this test. Technical Report #1 (Bliss, 1982) describes, in detail, the process used to choose an appropriate test. Finally, the Stanford Achievement Test (1973 version) was chosen as the instrument of choice. Briefly, the reasons for choosing this achievement test battery were that 1) it covered the grades K-12, 2) it seemed, on initial observation, to be a good match with subject content taught in the schools, 3) it was technically sound given the population on which it was standardized (a seemingly representative sample of mainland U•S• students), and 4) it would report out criterion referenced results.

Due to various organizational and fiscal contraints only students in even numbered grades were tested. This seemed acceptable since many of the objectives tested by the Stanford

Achievement Test carry across adjacent levels of the test and there was no reason to suspect that the patterns of academic achievement of students in odd numbered grades were different from those in even numbered grades. These constraints are detailed in Technical Report #1.

Sampling of students to be tested was done using a cluster sampling technique with classes as the cluster unit. The details of the sampling procedure can be found in Technical Report #1 along with a discussion of the effects of using this sampling technique on the information obtained. Table 1 presents a breakdown of the sample as it finally emerged.

Table 1

U.S. Virgin Islands Sample Sizes

Grade	Test Level	Total System	St.Thomas St. John District	St. Croix District
12	TASK II	129	74	55
10	TASK I	254	167	87
8	Advanced	345	173	172
6	Intermediate II	227	146	81
4	Primary III	346	186	160
2	Primary I	234	143	91
	Total	1535	889	646

Testing was done at the grade level recommended by the test publisher during the week of October 21, 1980 in the St. Thomas/St. John district and during the week of December 1, 1980 in the St. Croix district. Testing materials and completed answer sheets were collected, answer sheets checked to determine compliance with marking instructions, and answer sheets sent off to be machine scored.

Background - Technical Report #2

This is the second of four reports that will deal with the results of the basic skills assessment described above. Technical Report #1 detailed the procedures used in test selection, sampling, and test administration. More importantly, it established empirically, the content validity and the reliability of the Stanford Achievement Test when it was administered to a sample of U.S. Virgin Islands public school students. This is particularly important since there exists no standardized test of academic achievement which includes USVI students in its standardization group.

This report examines the scores of secondary school students (grades 8, 10, and 12) and presents :

1) an item analysis of each item on each test of the battery which includes indices of item difficulty and discrimination and
2) a summary of student skills based on their scores on the items keyed to specific objectives.

Item Analysis

Difficulty indices

The difficulty index of an item is found by taking the number of examinees scoring correctly on an item and dividing it by the total number of students taking the test. In short, it is the proportion of examinees who scored correctly on an item and has a range from 0 (no students scoring correctly) to 1 (all students scoring correctly). Because of the fact it is a proportion, it is often designated in the literature as "p" (e.g. p=.75). It may be worthwhile to point out that the term "difficulty index" may be somewhat of a misnomer since items with high difficulty indices are actually less difficult than items with low difficulty indices. Nevertheless, the term and its definition have become standard in the area of psychometrics throughout the United States.

Difficulty indices for each item on each test were reported out by the test scoring service. In addition, difficulty indices for examinees in the standardization group in the same grade as local examinees at approximately the same time of year are reported. The test scoring service used a Chi-squared test for proportions on each difficulty index to test the hypotheses that the proportions of local students scoring correctly on individual items in greater or less than the proportions of examinees in the standardization group scoring correctly at the .05 level of significance. Significant differences in either direction were reported out.

Discrimination Indices

The item discrimination index indicates the degree to which responses on one item are related to responses on other items on the test. The statistic indicates whether a person who does well on the test as a whole (that is, a person who is presumably high on the trait being measured) is more likely to get the particular item correct than a person who does poorly on the test as a whole. In other words, the item discrimination index indicates whether an item discriminates between those who do well and those who do poorly on the test as a whole. Taking the item difficulty and the item discrimination index into consideration, the developers of tests desire to construct tests which discriminate well among examinees with varying levels of a trait.

The item discrimination index is calculated by the formula

$$d = \frac{U - L}{N}$$

where

U = the number of examinees who have total test scores in the upper range of total test scores and who also have the item correct.

L = the number of examinees who have total test scores in the lower range of total test scores and who also have the item correct.

N = the number of examinees in the upper or low range of the test scores.

By definition, d is the difference between the proportion of high scoring examinees who got the item correct and the proportion of low scoring examinees who got the item correct. Upper and lower ranges generally are defined as the upper and lower 10% to 33% of the sample, with examinees ordered on the basis of their total test score. When total scores are normally distributed, using the upper and lower 27% produces the best estimate of d (Kelly, 1939). If the distribution of total test scores is flatter than the normal curve, the optimum percentage is larger and approaches 33% (Cureton, 1957). However, Aller and Yen (1979) found that, for most applications, any percentage between 25 and 33 will yield similar estimates of d. In this study, 27% was used as the upper and lower percentage because examination of selected distributions of actual test scores revealed nearly normal distributions.

The theoretical range of d is between -1 and +1. However, maximum discrimination is likely to occur when p=.50. When p=.50 the variance in item scores, which is p(1-p), is maximized. As an item becomes more difficult, it is less likely that <u>any</u> student will score correctly on it. As it becomes less difficult it is more likely that <u>any</u> student will get it correct. This could lead to the suggestion that all items should have p=.50, but the usefulness of this suggestion is mitigated by interrcorrelations among items. In an extreme case, if the items on a test all interrcorrelated perfectly and had difficulties of .50, half the examinees would receive a total test grade of zero and the other half would have

perfect test scores. Hence, there would be no fine discrimination between examinees' levels of achievement or whatever trait was being measured. In general, test designers tend to try to choose items with a range of difficulties that average around .50. Items of particularly low difficulty are often included in tests (usually among the earlier items) for motivational reasons.

Item discrimination indices were calculated in this study to provide indications that items may be flawed when used with USVI students. Such flaws are ambiguity, the presence of clues, the presence of more than one correct answer, and other technical defects. If none was found upon examination of the item, and it was determined that the item did, indeed, appear to measure the objective it was intended to, the item was included in the overall analysis of results. Any item that discriminates positively can make a contribution to the measurement of pupil achievement and low indices of discrimination are frequently obtained for reasons other than item defects.

Standardized achievement tests are designed to measure several different types of learning outcomes (e.g. knowledge, understanding, application, etc.). Where this is the case, test items that represent an area receiving relatively little emphasis will tend to have poor discriminating power. For example, if a test has forty items measuring knowledge of specific facts and ten items measuring understanding, the latter items can be expected to have low discrimination indices. This

is because the items measuring understanding have less representation in the total test score and there is typically a low correlation between measures of knowledge and measures of understanding. Low discrimination indices here merely indicate that these items are measuring something different from what the major part of the test is measuring. Removing such items from the test would make it a more homogenous measure of knowledge outcomes, but it would also damage the content validity of the test because it would no longer measure objectives in the understanding area. Since achievement test batteries need to measure a wide variety of objectives in a reasonably short period of time, they tend to be fairly heterogeneous in nature and moderately low discrimination indices tend to be the rule rather than the exception.

To summarize, a low discrimination index alerts test users to the possible presence of defects in test items but does not cause them to discard these items if they appear to be functioning as they should. A well constructed achievement test will of necessity, contain items with low discriminating power and to discrard them would result in a test which is less, rather than more, valid. Due to these considerations, in this study items were examined if they had discrimination indices lower than .20. This is a rather conservative criterion since items that discriminate as low as this may provide useful information, but given the unknown test taking characteristics of USVI students, it was decided to be particularly cautious in the item analysis.

Summary of Student Skills

The items on the tests of the various levels of the Stanford Achievement Test battery were keyed to behaviorally stated instructional objectives. These objectives are grouped into a series of item groups. Tables are available which present objectives within item groups and the difficulty and discrimination indices for the item or items which evaluate those objectives. These may be obtained from the Caribbean Research Institute. In addition, the difficulty index for each item for the examinees in the standardization group is available. This standardization group consists of examinees who were in the same grade at approximately the same time during the school year as the U,S, Virgin Islands sample.

The national p values are used not as a means to compare U•S• Virgin Islands students with mainland U•S• examinees. Historical and cultural differences between these two groups of examinees makes this comparison an inappropriate one. Philosophical considerations aside, however, such comparisons are of little use to the people who make curricular decisions in schools. What these people need to know are the particular levels of skills of students as measured against well defined objectives, not how well their students achieved thse skills as compared to other students.

Nevertheless, since the skills and knowledges taught in schools are seldom taught once, but are dealt with at a number of grade levels where they are reinforced and broadened, the level of achievement on specific objectives should be expected

to change from grade to grade for a particular student or group of students. The publishers of the Stanford Achievement Test take this factor into consideration by testing particular objectives across a number of levels (grades) of the test. In this study the national p values are used to indicate the relative level of difficulty of an item by which the performance of the local sample may be judged. For instance, it would be foolish to be dissatistified if 20% of the USVI students have indicated that they have reached an objective when only 18% of comparable students in the standardization group had reached that same objective. What is more likely the case is that this is a difficult and complex objective that had just been taught recently and would be retaught and elarged upon at a later time.

Therefore, the following criteria were used in summarizing student skills. Skills are described as "adequate" if the proportions of local examinees scoring correctly on items measuring those skills are not significantly higher or lower than the proportions of the standardization group scoring correctly or, if significantly higher or lower, the proportions correct are within 10% of the standardization group proportion correct as reported by the test scoring service. Skills are described as "strong" if the proportions of local examinees scoring correctly on items measuring those skills are significantly higher and more than 10% greater than the standardization group proportion correct as reported by the test scoring service. skills are described as "weak" if the proportion of local examinees scoring correctly on items measuring those skills are

significantly lower and more than 10% less than the standization group proportion correct. The 10% standard was set in the realization that some differences, while statistically significant may be educationally trivial and it was noted that most differences indicated as significant exceeded 10%.

Finally, in cases where the scores of examinees from both the St. Thomas/St. John and the St. Croix districts were the same based on the criteria stated above summaries were based on the entire USVI sample. Where differences were noted, the skills of examinees within each district were summarized separately.

Availablility of Data

Tables listing each objective on every test keyed to test items with national difficulty indices, difficulty indices for the entire USVI sample and the individual district subsamples as well as item discrimination indices for the whole sample and each subsample are available and may be obtained from the Caribbean Research Institute, College of the Virgin Islands.

Grade 8 - Advanced Level

The Advanced Level of the Stanford Achievement Test Basic Battery consists of seven tests with the number of items listed below:

1. Vocabulary - 50 items
2. Reading Comprehension - 74 items
3. Mathematics concepts - 35 items
4. Mathematics Computation - 45 items
5. Mathematics Applications - 40 items
6. Spelling - 60 items
7. Language - 70 items

The means, standard deviations, and reliability estimates for these tests determined previously (see Bliss, 1982) are presented below:

Test	Mean Raw Score	Standard Deviation	KR-20 Reliability	Standard Error
		Total USVI Sample		
Vocabulary	21.0	7.2	.81	3.15
Reading Comprehension	31.5	15.4	.95	3.45
Math Concepts	15.3	5.8	.74	2.65
Math Computation	23.0	7.6	.85	2.95
Math Application	16.9	6.6	.83	2.72
Spelling	31.8	12.3	.93	3.27
Language	35.6	12.1	.88	4.20

		St. Thomas/St. John Sample		
Vocabulary	20.7	6.6	.78	3.10
Reading Comprehension	32.6	17.4	.97	3.02
Math Concepts	16.4	6.0	.81	2.61
Math Computation	23.3	7.1	.83	2.92
Math Applications	17.7	6.6	.83	2.72
Spelling	32.8	11.8	.92	3.33
Language	35.8	10.6	.84	4.23
		St. Croix Sample		
Vocabulary	21.2	8.5	.87	3.05
Reading Comprehension	30.5	13.1	.93	3.46
Math Concept	14.2	5.3	.75	2.67
Math Computation	22.8	8.2	.87	2.94
Math Applications	16.1	6.5	.82	2.76
Spelling	30.8	12.9	.93	3.40
Language	34.6	13.6	.91	4.07

Following are the item analyses and summaries of student achievement on the objectives evaluated by each of the tests.

Vocabulary

Item Analysis

All items showed acceptable discrimination indices with the following exceptions:

Item #16 - This item was an easy one (p=.72) for USVI examinees regardless of overall achievement.

Item #32 - This item was particularly difficult for USVI students (p=.16). It called for them to select a synonymn for "ponder." Many students chose "beat" over the correct response "think."

Item #39 - This item was particularly difficult (p=.17), but there seemed to be no particular pattern of incorrect responses.

Item #40 - This item simply did not discriminate well between high and low achievers, although difficulty was moderate (p=.37). It required students to select a synonym for "contemplative" and many examinees chose the distractor "cautious" over the correct response "thoughtful."

Item #43 - This item discriminated very poorly between high and low achievers. High achievers often choose "dejected" as a synonym for "exhilarated" rather than the correct response "stimulated."

Item #47 - This item discriminated very poorly in spite of a moderate difficulty index (p-.29) and there appeared to be no pattern of incorrect responses.

Item # 29 - This item showed very little discrimination with high achievers equally likely to answer the item correctly as were low achievers with a moderate difficulty index (p=.52).

Item # 34 - This item showed low discrimination between high and low achievers. The difficulty index (p=.31) was not particularly low nor was there a noticeable pattern of incorrect responses,

Item # 1 - This item was very easy for USVI students (p=.96).

Item # 30 - This item tended to be difficult for USVI students (p=.21) with no patterns observed in incorrect responses.

Item # 31 - This item was a difficult one for USVI students (p=.21).

Item # 44 - This item discriminated satisfactorily for St. Croix school district students, but poorly for St. Thomas/St. John district examinees. These latter examinees in the high achieving group chose "dangerous" very often as a synonym for "anomolous" over the correct response. This pattern was not evident among St. Croix school district students.

Item # 41 - This item discriminated satisfactorily for St. Croix district students, but did not discriminate at all among high and low achievers in the St. Thomas/St. John district. These examinees tended to choose "active" as a synonym for "audacious" over the correct response, "bold." No such pattern was noticeable among St. Croix district students. In general, this tended to be a difficult item for students in both St. Thomas/St. John and St. Croix districts (p=.19 and .21,respectively).

Item # 1[illegible] - This moderately difficult (p=.40) item tended to discriminate poorly among high and low achievers. There appeared to be no ambiguity in the item and no observable pattern of incorrect response.

Summary of Student Skills

In general, USVI students tested seem to be weak in the area of standard English vocabulary recognition. Most items seem to function well with most poorly discriminating items being the result of either extremely high or low levels of difficulty. Since few of the incorrect response patterns (i.e. examinees choosing a particular incorrect item with a high frequency) which typically mark ambiguous items were present, it is unlikely that these scores were the result of technical difficulties in the item, themselves. Of the fifty items on the test, a significantly smaller proportion of USVI students were able to respond correctly when compared to the standardization group on 39 of these items. Beyond the question of statistical signigicance, the proportions of USVI students who successfully answered many items was substantially lower than the the standardization sample. The differences in item mean difficulty values are a clear indication of this problem.

Of particular note is the single item on which the USVI sample outperformed the standardization sample (but only in the St. Thomas/St. John district). This item asked examinees to choose the synonym for "capsized" by stating, "A capsized boat has been ---" with the correct response choice being "overturned." Since such an item may allude directly to personal experiences of USVI students, this is a phenomenon that bears examining in looking at other items.

Reading Comprehension

Item Analysis

All items showed satisfactory discrimination indices except the following:

Item # 12 - High and low achieving students seemed to be equally likely to correctly answer this moderately difficult (p=.30) question in the St. Croix district. In the St. Thomas district, low achievers appeared more likely to be able to correctly answer the item than were high achievers. High achieving students seemed to choose distractor # 4 very often. The item requires examinees to choose the most appropriate title for a passage they have read. Distractor # 4 refers to a particular section of the passage and uses names and ideas in that section. However, it does not reflect the overall idea of the passsage. Students choosing this distractor are not aware of this overall (or "global") idea.

Item # 54 - This item seemed to discriminate satisfactorily for examinees in the St. Croix district, but less well for those in the St. Thomas/St. John district. This was a fairly difficult question (p=.18) which required examinees to choose the main idea in a short poem by Walt Whitman. A large number of students from both districts seem to have omitted this item which appeared toward the end of a 74 item examination. This will be commented on in the summary. Item # 54, 60, and 74 display the same phenomena as item #54, showing many omits on what seem to be difficult items.

Item # 3 - This appeared to be an easy item for most students (p=.90)

Item #57 - This item appeared to be a difficult one for students in

both districts (p=.18). However, it should be noted that 48% of students tested omitted this item and this could account for the low difficulty and discrimination indeces.

Item # 73 - This item seemed to exhibit the same phenomenon as item # 57. In this case 59% of students tested omitted this item.

Item # 41 - In this fairly difficult item (p=.28) high achieving students were distracted by choice # 4. In this question students read a paragraph about home mortgages and were asked to choose a statement indicating the value of a mortgage to a lender although this advantage was not explicitly stated. Choice # 4 indicated that the value was that the lender would "eventually own the property," which the paragraph indicated was the consequence to default on the part of the borrower. The choice which was considered correct was, "his money earns interest," and the fact that the borrower pays interest was also explicitly stated in the passage. Since either of these situations could be considered as "value" by a thoughtful examinee, it may be that the question was defective in that it had two answers that could be considered correct by students who lacked experience with mortgages.

Item # 72 - This item appeared to be very difficult for USVI examinees. However, 59% of examinees omitted the item and this could account for the low difficulty and discrimination indices which were observed in the St. Thomas/St. John examinees responses.

Item # 65, 67, 69 - These three items showed satisfactory discrimination in the St. Croix district, but poor discrimination in the

St. Thomas/St. John district. One phenomenon that was noted was that while the numbers of high achievers scoring correctly on these items was fairly constant between districts, the number of low achievers scoring correctly was greater in St. Thomas/St. John than in St. Croix. Also, the proportion of omits was higher in the St. Thomas/St. John district (63,64, and 66%, respectively vs 52,51, and 52% on St. Croix) and this could account for the differences in item discrimination power that was observed.

Item # 23 - This moderately difficult item (p=.34) discriminates little, if at all, between high and low achieving students. In general, low achieving students seem just as likely to obtain correct scores on this item as high achieving students. There does, however, seem to be differences in the patterns of incorrect responses between these two groups of examinees. The item asked students who had read a passage which dealt with animals supported by the British government which included a cat who was the "Official Rat Catcher" in the storeage cellars of London's government offices. The passage indicated that the cats food cost the government the equivalent of 65¢ per week while neglecting to give any information on the cost of upkeep for any of the other animals in the passage, the question, "the writer probably mentions exactly what it costs to feed Peta [the cat] because he thinks it is ---," should elicit "amusing" as the correct response. Most examinees failed to see this as being amusing, with the low achievers tending to select "too low" and the higher achievers tending

to select "unfair" as their answers. The latter seems to be more of a moral judgement and it may be significant that this was the choice of the higher achieving students. It is also interesting to note that the Technical Data Report published by the testmaker shows this item to discriminate less well than most other items (although much better than on the USVI sample of examinees) when the test was administered to the standardization groups.

Item # 29 - This moderately difficult item (p=.36) discriminated poorly between high and low achievers. It required students to infer something about the climate of Tropical Africa from reading a passage concerning rivers in the region and their great potential for the production of hydroelectrical power. Most students concluded that in this region there was "sunshine most of the time," rather than the conclusion drawn from the passage that there was "quite a bit of rainfall." What may be happening is that examinees are drawing from their own previously learned concepts about the nature of the African climate and/or their personal experiences concerning the climate in a tropical zone instead of using the information provided in the written passage they were to read.

Items # 58, 61, 63, 66 - These items appeared to be very difficult for USVI examinees (p=.15, .14, .17, and .12, respectively). However, as was the case in other items toward the end of the test, the number of students omitting these items was very high (52, 56, 59, and 63%, respectively in the St. Thomas/St. John district and 49, 52, 54, and 56% in the St. Croix

district). These high proportions of omitted responses could account for the low difficulty and discrimination indices which were obtained.

Summary of Student Skills

Table 2 shows the proportions (in percents) of omitted responses on each item in the Reading Comprehension test. From this table it is apparent that students in this sample did not have the opportunity to attempt all of the items on the test in the time allowed to them. This is an important observation. If students could not attempt items, they could not possibly score correctly on them and this would tend to deflate the difficulty indices of items toward the end of the test, making scores of USVI students seem lower than they should actually be. In addition, since less than half the students were able to attempt items 60 through 74, we would expect lower discrimination indices than would be obtained had all students been able to complete the test. This phenomenon seems to be unique to the Reading Comprehension test. On all other tests in the Advanced Level battery examinees appear to have ample time to finish within the time provided by the test publishers. Any attempt to explain why the USVI group required more time than the standardization groups would only be speculation, but further research should be done to investigate the possibilities that such factors as a more deliberate reading style (resulting in longer periods of time being required for students to read each passage), fatigue, and differences in attention span might be contributing to this phenomenon.

Given these observations, it might be well to be cautious of

Table 2

Proportions of the Eighth Grade Sample
Omitting Reading Comprehension Test Items
(Percents)

ITEM#	USVI	STT/STJ	STX
1	1	1	0
2	0	0	0
3	0	0	0
4	1	1	1
5	1	2	1
6	0	0	0
7	1	1	1
8	1	1	0
9	0	0	0
10	1	1	1
11	1	2	1
12	1	1	1
13	0	1	2
14	1	2	1
15	1	1	1
16	1	1	1
17	0	1	0
18	1	1	0
19	1	2	1
20	0	1	0
21	1	1	1

Table 2 Continued

22	1	2	1
23	0	1	0
24	1	2	0
25	4	5	3
26	4	5	2
27	4	4	4
28	4	5	3
29	4	5	3
30	5	5	5
31	6	5	6
32	8	8	8
33	8	8	9
34	7	6	8
35	8	8	8
36	9	7	11
37	11	9	13
38	13	11	14
39	14	13	15
40	14	12	15
41	15	14	17
42	18	17	19
43	19	19	19
44	20	20	20
45	21	20	21
46	23	23	23
47	29	31	27
48	30	31	28

Table 2 Continued

49	32	34	30
50	34	38	30
51	35	38	31
52	39	44	34
53	41	45	37
54	42	45	40
55	44	48	40
56	47	49	49
57	48	52	44
58	49	52	46
59	50	53	48
60	51	55	48
61	52	56	47
62	53	58	48
63	54	59	50
64	55	60	51
65	57	63	52
66	56	63	50
67	58	64	51
68	57	64	51
69	59	66	52
70	59	67	51
71	58	65	52
72	59	63	54
73	59	64	59
74	58	63	53

drawing any conclusions from the results of items 50-74 where more than a third of the examinees omitted the items in question.

Restricting the observations to items 1 through 49, however, the following generalizations can be made.

USVI students tested seem to be somewhat weak in the area of determining the general idea of the passages they read, although this must be a somewhat tenuous judgement based upon the small number of items (five) from which this judgement could be made because of the large number of examinees omitting later items of the test. However, in three of these five items, local students had a significantly smaller proportion of correct answers than did the students in the standardization groups.

In the area of determining meanings explicitly stated in a reading passage, USVI students appeared to have achieved well. In most of the items which most students seem to have had the opportunity to answer, the proportion of local students answering correctly was as high as or only slightly lower than the standardization group's proportions correct.

Implicit meaning refers to ideas in a piece of reading which are not specifically in the reading, but must be deduced from facts or other information presented in the test. The USVI examinees appear to have difficulty in determining explict meanings in passages they have read. In ten out of thirteen items dealing with this skill that most students were able to attempt to answer in the time allowed, the students in the local sample scored correctly a significantly lower proportion of the time than standardization sample. Further, the mean proportion of students

scoring correctly on these items was particularly low (59% vs 70% for the standardization group).

Often a reader must obtain information from a text by obtaining clues from what is explicitly stated within the text. The USVI examinees had great difficulty in determining meaning from written passages on the basis of contextual clues. In seven out of the nine items that most examinees were able to attempt within the time allowed, the proportion of local students answering items correctly was significantly lower than the proportion of examinees in the standardization sample answering the same items correctly. Again, these proportions were particularly low with a mean of 38% for the local examinees versus a mean of 52% for the standardization group.

The students in the USVI sample had great difficulty in tasks which required them to make inferences or draw conclusions from passages they had read. The proportion of USVI examinees answering items correctly was significantly lower than the proportion of examinees in the standardization sample who answered the same items correctly for all of the eleven items which most examinees in the USVI sample were able to attempt in the alloted time. Determining inferential meaning appeared to be the reading comprehension skill which gave USVI students the most difficulty with a mean correct proportion on the above mentioned eleven items of 28% versus 46% for the standardization group.

An observation that may be significant is that USVI students showed strength in the area of determing explicit meaning from material which they read while showing weaknesses in achievement

in other areas of reading comprehension. The taxonomy of educational objectives developed by Bloom (1956) may be a good theoretical framework from which to view this phenomenon. Being able to derive meanings that are explicit in readings involves objectives at the comprehension level of the taxonomy while the other objectives types (deriving global meaning, implicit meaning, inferential meaning, and meaning from context) involve higher level (and, therefore, more difficult) objectives in the area of analysis and synthesis. It appears that the sample of USVI students may have, as a group, achieved reading comprehension objectives at the comprehension level, but have not, as yet, fully mastered many of the higher level objectives in this skill area.

Mathematics Concepts

Item Analysis

All items discriminated satisfactorily except the following:

Item 6 - This was a particularly difficult item for USVI students (p=.15). One distractor operated to cause many students to choose it over the correct response. Specifically, the item asked students to complete the arithmetic sentence: 6÷12=. Most students, regardless of general level of achievement, chose "2" as the correct answer rather than "½". This indicates that the item was functioning properly, but that examinees had not mastered the objective being tested.

Item # 4 - This proved to be a relatively easy item for examinees (p=.75), but it did not discriminate satisfactorily in the St. Thomas/St. John district. This was primarily because low achieving students were almost as likely to answer correctly as were high achieving students.

Item # 16 - On this moderately difficult item (p=.36) high achieving students were not much more likely to answer correctly than were low achieving students. There appeared to be no pattern observable in the incorrect responses.

Item # 29 - In this moderately difficult item (p=.35) high achieving students were not much more likely to answer correctly than low achieving students. The item asked students to determine the effect on a factor if the quotient remained constant and the other factor decreased. Many students from both achievement level groups chose the response that indicated that the factor decreased rather than the correct response indicating that it would increase. The item appears to be operating properly, however, since this error is a common one and indicates a lack of understanding of the concept.

Item # 33 - On this difficult item (p=.27) most students in both achievement level groups scored incorrectly. The lack of discrimination was very likely due to the difficulty of the item rather than to any technical fault in its writing. No response patterns were observed in the incorrect answers.

Item # 35 - This was an extremely difficult item for all students (p=.22) resulting in a low level of discrimination. There

appeared to be no pattern of incorrect responses.

Summary of Student Skills

Number

The item grouping "Number" refers to the concepts which deal with the relationships between different types of numbers (e.g. primes, odds vs evens, fractions vs whole). USVI students showed no particular strengths in this area. Concept areas that appeared to be particularly weak were:

a) the concept of fraction and its relationship to other types of numbers.

b) the concept of odd and even numbers

c) the concept of signed numbers

d) the concept of "modulo"

Notation

"Notation" refers to the technique by which numbers are expressed in written form. Such forms include expanded notation, differing base values, decimals, and fractions. The students in the USVI sample seemed to have a strong grasp of the concept of exponential notation. They seemed to be weakest in the following areas:

a) use of expanded notation

b) decimal notation

c) translating arabic numbers into word form

Operations

The term "operations" refers to the manipulation of numbers based on the use of a set of previously defined functions (e.g. addition, subtraction, multiplication). The students in the local

sample showed strenghts in the area of multiplication of decimal fractions, but seemed to be weak in the following areas:

a) solving inequalities

b) identifying multiplicative identity

c) identifying the uses of the associative property

d) the use of signed numbers in multiplication and division

e) identifying formulas for qualities of geometric functions

f) estimation in multiplication and division

Geometry and Measurement

This appeared to be the area of math concepts in which the USVI sample exhibited strongest levels of achievement. Examinees performed well in the following areas:

a) metric conversion

b) the use of Venn diagrams

c) graphing

Areas in which weaknesses were observed were:

a) set theory

b) estimation of angle sizes

c) properties of geometric figures

d) probability

Mathematics Computation

Item Analysis

All items discriminated satisfactorily with the exception of the following.

Item # 36 - This item appeared to be very difficult for examinees in the USVI sample (p=.10). The question required students to find the average of 12 and 6. In the low achieving group there appeared to be no pattern of incorrect responses, but high achievers were noted to choose the response "18" a large proportion of the time. This response would occur if examinees summed to two numbers, but failed to follow through on the second part of the averaging operation and did not divide.

Item # 45 - This appeared to be a difficult item for the examinees in the local sample (p=.24). It should be noted that the item was also difficult and discriminated at low levels (d=.29) for the standardization sample. An examination of the item, however, indicates that it is quite clearly worded and the task seems to be simply presented. The lack of discriminating power is most likely due to the general difficulty of the item.

Summary of Student Skills

Knowledge of Primary Facts and Solution of Simple Mathematical Sentences

On these skills there appears to be some substantial differences between the achievement levels of students in the St. Thomas/St.John and the St. Croix district. Because of this it would be useful to summarize achievement characteristics in each district, individually.

In the St. Thomas/St.John district students showed strong skills in the following areas:

a) use of the associative property of multiplication

b) multiplication and division with integers, fractions, and mixed numbers

c) addition, subtraction, and renaming of fractions

d) use of the distributive property

Weaknesses were observed in these areas:

a) multiplication with carry facts

b) operations with negative terms

c) renaming exponential numerals

In the St. Croix district examinees exhibited strong achievement in these areas:

a) multiplication and division of integers, fractions and mixed numbers

b) use of the distributive property

c) appropriate use of the distributive property

Weaknesses in the following areas were observed:

a) use of the associative property in multiplication

b) operations involving negative terms

c) multiplication with carry facts

Addition and Subtraction Algorithms

Objectives in this item grouping deal with the mechanics of addition and substraction. Students in the USVI sample showed good achievement in this area on all objectives.

Multiplication and Division Algorithms

This item grouping tested objectives dealing with the mechanics of multiplication and division. Students showed good achievement levels on all objectives in this area with the exception of the ability to find the quotient in a division problem with a

two digit divisor and a zero in the ones column of the quotient. The majority of students missing the item which tested this objective indicated "NH" (not here) as their choice leading the researchers to suspect that they came up with an answer through faulty use of the algorithm.

Common Fractions

St. Thomas/St. John district students showed particular strengths in finding common denominators and in working problems where common denominators must be found. They achieved adequately on the objective of finding the common fraction of a whole number. However, they appeared weak in the area of fraction multiplication.

St. Croix district students appear to have achieved adequately in the area of fraction multiplication, but were weak on objectives which required them to find and work with common denominators.

Students in both school districts appeared to have particular difficulty in the area of division by fractions.

Other Operational Models

All USVI students achieved strongly in the objectives involving exponential notation and prime factorization. Their achievement seemed adequate on objectives dealing with decimal notation and number lines. Weak areas were on objectives having to do with finding averages and working with percentages.

St. Thomas/St. John district students exhibited good achievement levels in all other areas.

St. Croix students showed good achievement levels in other areas with the exception of objectives dealing with common and simple algebraic manipulation.

Mathematics Applications

Item Analysis

All items discriminated adequately except the following.

Item # 11 - This was a particularly difficult item for USVI students (p=.20). High achieving students were only slightly more likely to score correctly on this item than were low achieving students. There appeared to be no pattern in the distribution of incorrect answers, except that choice number 4 (which indicated that a shaded area of a square took up the entire square when it obviously did not) was chosen very seldomly.

Item # 35 - This appeared to be a difficult item for the local sample students (p=.29). It was the only problem dealing with metric units on the test. However, the metric equivalents were given in the stem of the item, so a student knowing nothing of the metric system could have determined the correct answer. A large number of students missing this item either left in blank or indicated the choice "NH" (not here). It is possible that these students were intimidated by the metric units and did not even attempt the item and that this phenomenon carried through both high and low

achieving students.

Item # 37 - This item discriminated poorly. It tended to be a difficult item (p=.21) and asked students to determine the sum of six angles marked off in triangle with a perpendicular drawn from one of the angles to the opposite side. The student selecting the proper answer needed to put together the facts that the sum of the angles of a triangle is 180^{o} and that a perpendicular drawn to a side produces two angles of 90^{o} . The incorrect choice chosen most frequently was "180^{o}" which indicated that examinees may have been aware of this first fact, but were unaware of the second or were unable to use the two facts together.

Item # 40 - This was a very difficult item for USVI students (p=.14). This less than chance proportion of students scoring correctly on this item and the large proportion of omits (20%) evokes the possibility that, although it is part of the curriculum, the objective was never really taught in the schools, and/or that being the last item on the test, a substantial number of students never had the chance to attempt to answer it.

Summary of Student Skills

Analysis and Development of a Solution

Design, Selection of a Solution Sentence

and Adequacy of Data

St. Thomas/St. John district examinees exhibited adequate skills in solving problems which required them to:

a) divide to determine shares

b) add and divide in multistep sequences

c) translate verbal problems into mathematical sentences

d) determine correct change in a two step sequence

They displayed weaknesses on items requiring them to:

a) determine the adequacy of given data for solving a problem

b) express verbally the process for solving a problem

c) estimate the ratio of two numbers greater than 1,000

St. Croix examinees showed adequate achievement on none of the objectives in this item grouping. They showed particular weaknesses on items requiring them to:

a) divide to determine shares

b) translate verbal problems into mathematical sentences

c) determine the adequacy of given data for solving a problem

d) determine correct change in a two step sequence

e) express verbally the process for solving a problem

f) estimate the ratio of two numbers greater than 1,000

Date, Scale, and Percent

The USVI sample showed adequate achievement only on

the objective of determining a specific rate when one rate is expressed as an average. Particularly marked weaknesses were found on items which required the examinee to:

a) express part of a total as a fraction

b) determine specific rate when no average rate is given

c) interpret map scales

d) find a total when a part and its fractional representation are known

e) determine percents

Measurement

Examinees in the USVI sample demonstrated adequate achievement in solving problems which required them to:

a) find the volume of a rectangular solid

b) find the area of a rectangle

c) find the sum of six angles

They showed weaknesses in solving problems which required them to:

a) calculate the area of a shaded region

b) convert standard measures of weight in order to divide

c) add negative numbers

d) match English unit with metric approximations

e) find and compare areas of squares and parallelograms

f) identify finishing time when given starting and elapsed times

Students in the USVI sample demonstrated adequate achievement in identifying a function in a mathematical sentence with two variables and choosing the solution set for the sentence. They showed weaknesses in items dealing with the following objectives:

a) choosing the graph of a solution set

b) using line graphs

Statistics, Average, and Probability

St. Thomas/St. John district students demonstrated adequate achievement in the ability to compute an average. They showed great weakness in their ability to determine the probability of the occurance of a given event.

St. Croix examinees exhibited weaknesses across both of these objectives.

Spelling

Item Analysis

All items discriminated at a satisfactory level except the following.

Item # 10 - This was a very difficult item for USVI students (p=.18). Both high and low scoring students appeared to find it equally difficult.

Summary of Student Skills

Homophones

Homophones are words which sound alike, but have different meanings and are spelled differently. On these items,

students were presented with four homophones used in context and were required to choose the phrase in which the homophone was used inappropriately.

Students in the St. Thomas/St. John district showed particularly strong ability to recognize the distinction between the homophones "choral" and "coral" when used in context. They showed particular weaknesses in distinguishing missuse of the following homophones in context:

a) cereal/serial

b) straight/strait

c) niner/minor

d) arc/ark

On other homograph recognition pairs these students performed at an adequate level.

St. Croix district students showed particular weaknesses in distinguishing missuse of the following homophones in context:

a) cereal/serial

b) bored/board

c) wrap/rap

d) straight/strait

e) miner/minor

They showed adequate achievement in dealing with all other homophones in context.

Phonics

The students in the USVI sample showed a particularly strong grasp of the use of phonics in spelling. Specific objectives where local students achieved very highly were where

misspelled words had situations where

a) /le/was spelled e-e instead of ee

b) final /e/ was spelled ey instead of y

c) schwa sound was omitted

Objectives on which local students showed weaker achievement were tested by items where the misspelled words had situations where

a) /k/ was spelled c instead of ch

b) /kt/ was spelled ck instead of ct

c) /e/ was spelled i instead of e

d) /e/ was spelled e instead of ea

e) /ch/ was spelled d instead of t

f) /e/ was spelled e instead of a

It might be useful to note here that difficulties appeared to be of two types. First, in the use of vowels, particularly of those having the sound /e/. This sound is represented in a great number of ways in the English language, very often quite arbitrarily. As a result, it would be difficult to apply phonetic rules to its use. Second, in words that are generally mispronounced by the population at large. We tend to say "strickly" rather than "strictly" and "congradulations" instead of "congratulations." Again, applying phonetic rules to such words could cause difficulties.

Word Building

Word building skills refer to those which are involved in adding suffixes and prefixes to root words. USVI students showed very satisfactory achievement in this area. The few weak

areas if achievement are those in which students were required to identify misspelled words where

a) /ens/ is spelled ense instead of ence

b) a final l is omitted when adding ly

Language

Item Analysis

All items discriminated at a satisfactory level except the following.

Item # 16 - This item was an extremely difficult item for USVI students (p=.08). Low achieving students were just as likely to score correctly on it as high achieving studetns. Incorrect answers were spread in no particular pattern except that the incorrect answer was chosen less often than could be expected by chance (there were four choices). Only 2% of students tested omitted the item. However, it was noted that high achieving students responded in a particular way. The item required students to fill in the blanks in the sentence, "Beddoes soon saw that Humphrey was as good as ---; in fact, Humphrey was the --- of the two." The choices were 1) he-best; 2) him-better; 3) he-better; and 4) him-best, with choice # 3 listed as the correct answer. High achieving students tended to choose response 1 and 4 over the other two responses, indicating that they believed the superlative form of "good" was the corrcet response. Cooper (personal communication) points

out that persons who are considered to be exemplars of the use of formal standard English have adopted this usage. He suggests that it will become the formally accepted usage within ten years. Given this situation, the item may have been ambiguous since it could be thought of as having two correct answers.

Item # 6 - This item proved to be very difficult for students in the USVI sample (p=.16). It tested students' knowledge of both the capitalization conventions for school subjects and the use of commas to offset words in a series. Most students chose choices # 2 (Latin, Greek, Math and Science) and # 3 (latin, greek, math and science) with the former indicating they were not aware of the capitalization convention and the latter that they were unaware of both the capitalization convention and the comma usage. Choice #2 was, by far, the most popular. High achieving students who missed the item chose # 2 in almost every instance.

Item # 61 - This was a difficult item for examinees in the local sample (p=.23). There appeared to be no pattern of incorrect responses either for the high achieving or low achieving group. The item gave a dictionary entry for the word "versatile" including a pronounciation guide. Examinees had to be able to determine form the pronounciation guide that the last syllable of the word contained no vowel sound. In writing, and often in common pronounciation the word has an "i" in the

final syllable, but the letter is not sounded. What may have contributed to the difficulty here is that students attempted to answer the item without using the pronounciation guide as they were instructed to do and thus, tended to rely on their own pronounciation of the word because they could or would not make use of the pronounciation guide.

Summary of Student Skills

Capitalization Conventions

In the St.Thomas/St. John school district students seemed to have a good grasp of capitalization conventions tested except those used when naming school subjects. In this latter case they tended to capitalize the names of all school subjects.

Students in the St. Croix district had difficulty in dealing with the following capitalization conventions:

a) "Sir" as an honorific (e.g. Sir Walter Raleigh)

b) the name of a play

c) school subjects

d) names of religions

e) names of institutions and cities

On other capitalization conventions tested, these students appear to have achieved adequately.

Punctuation Conventions

Students in the St. Thomas/St. Croix school district seem to have difficulty in the use of commas in certain situations. They seem to be adequately familiar with comma usage under the following circumstances:

a) in dates

b) after an introductory participial phrase

c) in setting off appositive phrases

d) in setting off direct quotations

e) after introductory adverb clauses

f) before "and" when it joins two independent clauses.

g) between adjectives modifying the same noun

h) after introductory prepositional phrases

These students do seem to have difficulty in the use of commas under these circumstances.

a) setting off a parenthetical expression

b) when they are not used to set off indirect quotations

c) setting off direct quotations

d) setting off nonessential clauses

In the St. Croix school district examinees appeared to be adequately familiar with the use of commas in these situations:

a) in dates

b) in setting off appositive phrases

c) after introductory adverb clauses

d) when they are not used to set off indirect quotations

e) before "and" when it joins two independent clauses

These students seemed to show weaknesses in the use of commas under the following circumstances:

a) after an introductory participial phrase

b) setting off a parenthetical expression

c) separating items in a series

d) separating adjectives modifying the same noun

e) setting off direct quotations

f) setting off nonessential clauses

g) after introductory prepositional phrases

Usage Conventions

Usage conventions refers to the choice of particular forms of a word which are appropriate in a given context. Examples of this include the formation of past and present participles and the choice of appropriate verb tense.

Students in the USVI sample seemed to have a good grasp on the formation and appropriate use of the past participle of irregular verbs, but showed some weaknesses in determining the past participles of the verbs, "to write," "to begin," and "to wear." The latter verb only appeared to be troublesome to St. Thomas/St. John district students.

Examinees in both districts showed weaknesses in achievement on the following objectives:

a) choosing a verb form to avoid double negatives

b) choosing reflexive pronouns

c) choosing the appropriate verb when singular subjects are joined by "nor."

The three item groupings noted above (i.e. capitalization, punctuation, and usage conventions) form Part A of the language test and constitute the first 31 items of the test. This section is timed separately and it was noticed that after item #25, over 15% of the examinees tended to omit items. Twenty one percent of the examinees omitted item # 31. The items in this section

of the test require examinees to carefully read one or two sentence passage and choose the appropriate words to fill one or two blanks in the sentence or sentences. It was noted in the Reading Comprehension test that local students appear to take more time to read material presented to them with the result that they have difficulty getting to the later items on tests requiring much reading. It is possible that the same phenomenon is occurring on this section of the Language test, although on a smaller scale, with the result that a substantial number of examinees are not able to attempt the last few items on the test. This would have the effect of deflating difficulty indices and making the achievement of students in the local sample appear to be lower on the objectives tested by these items, and on the test as whole. It is suggested that the results obtained on these later items be viewed with caution when interpretations are made.

Distinguishing Fragments Sentences, and Run-On Sentences

Students in the St. Thomas/St. John school district appear to have little difficulty recognizing complete sentences and sentence fragments. On the other hand, they do seem to be weak in recognizing a run-on sentence when it is presented to them.

Students in the St. Croix school district appear to be able to recognize sentence fragments and run-on sentences, but seem to have difficulty recognizing complete sentences.

English Skills

This section of the test examines student achievement

in areas such as the use of dictionaries and other reference works, and in basic literary usage.

In the area of dictionary skills, students in the USVI sample showed adequate skills in using dictionary entries to determine the appropriate meanings of words used in context. They seemed to be weak in other areas of dictionary use such as use of the pronounciation guides and use of guide words.

Students in both districts had difficulty identifying the Reader's Guide to Periodical Literature as a source for finding the titles of magazine articles when the author was known and St. Croix district students exhibited difficulty in identifying an atlas as the reference one would use to find maps.

In the St. Thomas/St. John district students achieved adequately on the following objectives:

a) recognizing a word formed by a prefix or suffix

b) identifying meanings of morphemes

c) selecting the modal in the present tense

d) choosing sentences which have identical meanings

These students showed low achievement on these objectives:

a) identifying literary concepts

b) recognizing determiners that are used incorrectly

c) indicating reasons why given sentences are ungrammatical

d) identifying irrelevant words in sentences.

In the St. Croix district students achieved adequately on the following objectives:

a) identifying meanings of morphemes

These students showed weaknesses in achievement on the other objectives listed above.

Grade 10 - TASK I Level

The Stanford Achievement Test - Test of Academic Skills, Level I consists of three tests with the number of items listed below:

1. Reading - 78 items
2. English - 69 items
3. Mathematics - 48 items

The means, standard deviations, and the reliability estimates obtained for raw scores on these tests, as previously reported (see Bliss, 1982), are presented below

Test	Mean Raw Score	Standard Deviation	KR-20 Reliability	Standard Error
		Total USVI Sample		
Reading	45.6	14.0	.93	3.70
English	32.0	14.6	.98	2.07
Mathematics	48.0	12.0	.92	3.38
		St. Thomas/St. John Sample		
Reading	43.6	14.2	.94	3.48
English	32.1	.6.9	.99	1.69
Mathematics	47.6	10.8	.90	3.41
		St. Croix Sample		
Reading	48.5	14.3	.94	3.50
English	31.5	8.6	.89	2.85
Mathematics	49.0	14.0	.95	3.70

Following are the item analyses and summaries of student achievement on the objectives evaluated by each of the tests.

Reading

Item Analysis

Most items seem to discriminate at a satisfactory level. The exceptions to this statement will be examined shortly. Before doing this, however, it will be useful to note that, as in the Reading Comprehension test on the Advanced Level given to eighth grade students, large numbers of students seem to have been unable to attempt all the items on the test in the time allowed. Table 3 indicates the proportion of omits on each item in the Reading test. One difference from the eighth grade sample which is readily apparent is that the proportion of students omitting items in the St. Thomas/St. John district is considerably higher than that for the students in the St. Croix district. Also, the difficulty indices on items where a great number of omits are present seem to be higher for St. Croix students (i.e. more St. Croix students scored correctly on these items), whereas the proportions of students scoring correctly on items where there few omits do not differ greatly between districts.

From this it is clear that students in the St. Croix district are more likely to be able to attempt to answer later items on the Reading test than are students in the St. Thomas/St. John district and that this phenomenon results in generally higher scores on this test for St. Croix students than for St. Thomas/St. John students (mean raw scores are 48.5 and 43.6, respectively).

Reading

Item Analysis

Table 10 presents the difficulty and discrimination indices for items on the Reading test. Most items seem to discriminate at a satisfactory level. The exceptions to this statement will be examined shortly. Before doing this, however, it will be useful to note that, as in the Reading Comprehension test on the Advanced Level given to eighth grade students, large numbers of students seem to have been unable to attempt all the items on the test in the time allowed. Table 11 indicates the proportion of omits on each item in the Reading test. One difference from the eighth grade sample which is readily apparent is that the proportion of students omitting items in the St. Thomas/St. John district is considerably higher than that for the students in the St. Croix district. Also, the difficulty indices on items where a great number of omits are present seem to be higher for St. Croix students (i.e. more St. Croix students scored correctly on these items), whereas the proportions of students scoring correctly on items where there were few omits do not differ greatly between districts.

From this it is clear that students in the St. Croix district are more likely to be able to attempt to answer later items on the Reading test than are students in the St. Thomas/St. John district and that this phenomenon results in generally higher scores on this test for St. Croix students than for St. Thomas/St. John students (mean raw scores are 48.5 and 43.6, respectively).

Table 3

Proportions of the Tenth Grade Sample
Omitting Reading Test Items
(Percents)

ITEM #	USVI	STT/STJ	STX
1	0	1	0
2	0	1	0
3	0	0	0
4	1	1	1
5	0	0	1
6	4	4	3
7	2	2	1
8	0	1	0
9	1	1	1
10	1	2	0
11	1	1	1
12	0	1	0
13	4	5	1
14	0	1	0
15	1	1	2
16	1	1	0
17	2	2	1
18	2	2	1
19	2	2	1
20	2	2	0

Table 3 Continued

21	2	2	0
22	1	1	0
23	1	2	0
24	2	3	1
25	5	6	3
26	5	7	1
27	6	8	2
28	8	11	2
29	14	18	6
30	14	18	6
31	15	21	3
32	17	23	5
33	20	27	8
34	21	29	8
37	31	41	14
38	30	41	10
39	32	42	14
40	38	47	20
41	37	48	16
42	39	50	18
43	27	35	11
44	28	35	13
45	30	38	14
46	35	43	20
47	35	44	21
48	37	46	21
49	37	46	21

Table 3 Continued

50	39	48	23
51	41	49	26
52	5	7	1
53	9	11	6
54	8	10	6
55	8	10	5
56	9	10	9
57	14	14	15
58	9	9	8
59	14	14	14
60	14	14	14
61	18	19	16
62	13	14	13
63	18	19	16
64	20	21	17
65	17	19	14
66	20	21	15
68	19	19	18
69	21	22	20
70	20	22	16
71	23	25	20
72	24	26	21
73	25	28	20
74	24	27	20

75	28	31	22
76	30	34	22
77	30	34	22
78	33	38	24

Any attempt to explain these differences would be speculative at this time, but there are some possible explanations which come to mind. The most obvious is that St. Croix district students who took this test are more generally academically competent than their St. Thomas/St. John counterparts. This seem unlikely since other tests in the TASK battery do not exhibit this difference. Proportions of omits and raw score means are very similar (see Bliss, 1982).

Another possibility is that St. Croix counselors gave students time to finish the test and this violated the standard directions of the test. There is some evidence that this may have been the case. The examinees who were supervised by one counselor on St. Croix consistently left out fewer items toward the end of the examination than did the students supervised by the other counselor. In fact, the proportion, of students omitting items in the latter class was approximately the same (or slightly higher) than the proportion omitted by the entire St. Thomas/ St. John group of examinees.

Whatever the explanation, decision makers would do well to view the results of the Reading test as somewhat tentative, particularly on items where more than 1/3 of the students omitted the item. The following discussion will limit itself to those items which were answered by 2/3 or more of the examinees.

All items showed satisfactory discrimination indices except the following:

Items # 1, 3, 4, 22 - These proved to be very easy items for USVI students (p=.93, .95, .93, .94, respectively), and low achieving students were very likely to score correctly on these items.

Item # 64 - This item seemed to discriminate satisfactorily among St. Croix district examinees but poorly among St. Thomas/St. John examinees. The item required students to match the word "aspire" to the one word among five that was most closely related to it. This was a moderately difficult item for St. Thomas/St. John students (p=.26) and for some reason low achievers were just as likely to get the item correct in this district as were high achievers. No patterns of incorrect answers could be observed.

Summary of Student Skills

Global Meaning

Students in the USVI sample showed weaknesses in their ability to determine the main idea of a reading passage. In the St. Thomas/St. John district the proportion of students correctly responding to the four items with omit proportions under 33% was significantly and considerably lower than the proportions of the standardization group scoring correctly on all items (mean p for the St. Thomas/St. John group was .41 versus .55 for the standardization group).

In the St. Croix district the proportion of students correctly responding to the six items measuring global meaning was significantly and considerably lower than the standardization

group in 4 of the items. The mean p for the St. Croix group was .44 versus .56 for the standardization group.

Explicit Detail

The USVI sample of examinees appears to have achieved well in the area of determining explicit details from reading passages. On a third of the items (six out of eighteen) the proportion of local examinees scoring correctly was significantly higher than that of the standardization group. On only four of the eighteen items did local students score significantly below the standardization group.

Implicit Detail

Implicit meaning refers to information which is present in a body of text but is not explicitly stated.

Students in the St. Thomas/St. John district seem to do well on items requiring them to determine implicit details from a reading passage. Exceptions to this rule seem to be possibly related to the nature of the reading passage. For instance they tended to do poorly (p=.36 and .45 vs .60 and .59 for the standardization group) on items from a very long passage on sail boat safety. It is possible that reading passage length may be a factor here since other passages were considerably shorter.

St. Croix district examinees tended to score similarly to St. Thomas/St. John examinees with the exception of the fact that they seemed to have less trouble with one of the two items listed above where 52% of the sample scored correctly compared to 59% of the standardization group and only 45% of the St. Thomas/

St. John sample.

Inference and Logical Analysis

This section of the test evaluated students' abilities to draw logical conclusions from materials they have read and to make inferences concerning situations beyond the content of the readings. Students in the USVI sample appear to be adequately proficient in these skills with a mean group p value of .58 versus .55 for the standarization group.

Meaning from Context

These items appeared at the end of Part A of the test. Because of this, the proportions of omitted responses were extremely high indicating that many students did not have sufficient time to attempt to respond to these items (the mean proportion of omits for the entire sample on these items was over 34%). This phenomenon makes any attempt to interpret this data extremely suspect and, consequently, no attempt will be made here.

Word Meaning

This portion of the test evaluated standard English vocabulary.

In general, the USVI sample of examinees showed weak achievement in this area. St. Thomas/St. John district students' proportions correct were significantly lower than the corresponding proportions on the same items in all but one of the 27 items on this subtest. In many cases the differences were quite substantial with a mean p value of .51 for the St. Thomas/

St. John students compared to .69 for the standardization group.

In the St. Croix district examinees did somewhat better, but still showed weaknesses in vocabulary achievement. On 17 out of 27 items on this subtest the proportion of St. Croix district student correctly responding was significantly below that of the standardization group. Again, differences were quite substantial in many cases. The mean p value for the St. Croix district students on these items was .58 compared to a value of .69 for the standardization group.

English

Item Analysis

All items showed satisfactory discrimination except the following:

Item # 4 - This item discriminated satisfactorily between high and low achievers in the St. Croix district, but not in the St. Thomas/St. John district. This was a very easy item for students in both districts (p=.94 in both) with the result that most students correctly answered it. This accounts for the low discrimination values obtained.

Item # 12 - This item did not discriminate well on St. Croix. In fact, more students in the low achieving group scored correctly on it than students in the high achieving group. It is a moderately easy item (p=.80 on St. Croix). However, an examination of the incorrect response pattern failed

to show any reason for this phenomenon. The item asked the student to choose the situation when the term "love ya" would not be appropriately used. It provided the following choices:

1) in conversation with a friend
2) in conversation with an older person
3) in a letter to a friend
4) in a letter of application for a job

Choice # 4 was considered the correct answer. It was noted that high achieving students who responded incorrectly were no more likely to choose one distractor over another.

Item # 30 - This item discriminated poorly for examinees in the St. Croix district. It was a very easy item for these students (p=.94) and this should account for the low discrimination index obtained.

Item # 48 - This relatively easy item (p=.87) did not discriminate satisfactorily between high and low achieving students in the St. Croix district. From the three word, "askt", "trouble", "section", and "exspect" examinees were asked to choose the number of words spelled correctly. Low achieving students who scored incorrectly on the item showed no incorrect response pattern, but high achieving students who missed this item <u>all</u> chose "3" as the number of correctly spelled words. It seems that there is one of these two misspelled words that was not obviously misspelled even to students who spell well.

Items #53 and 54 - were very easy for St. Croix students (p=.97 and .90, respectively) and this should account for the low discrimination levels observed for these items.

Summary of Student Skills

Learning Skills

St. Thomas/St. John school district examinees showed weaknesses in the following areas of dictionary usage skills:

a) identifying a word that contains a given vowel sound

b) the change in principal stress in two forms of a word

They appeared to have achieved adequately on the other dictionary skills which were tested.

St. Croix examinees achieved adequately on all dictionary use skills tested except the ability to identify the change in the principal stress in two forms of a word.

While USVI students seem to be able to identify a reference to aid vocabulary development, they seem to have great difficulty in identifying a reference source for articles in periodicals.

The students in the USVI sample showed satisfactory achievement in objectives dealing with appropriate English usage. They were particularly strong in objectives having to do with identifying the appropriate usage of slang expressions and with identifying the action words in a sentence.

Usage Conventions

Examinees in the USVI sample showed weakness in their knowledge of the correct capitalization conventions for seasons

and common nouns.

St. Croix school district students exhibited adequate achievement in the choosing of appropriate comma usage, but St. Thomas/St. John examinees showed weaknesses in the proper use of commas between items in a series and of the use of commas after introductory words.

St. Thomas/St. John examinees showed some weaknesses on items which measured the appropriate use of verb tenses, but showed particular strength in dealing with objectives concerning proper use of pronouns and adjectives.

St. Croix school district students showed adequate achievement in all the areas listed above.

Spelling

USVI students showed particular strength in spelling. On all but two of the 15 spelling items, the proportion of USVI students scoring correctly was not significantly different or was significantly higher than that of the standardization group. The mean p value for spelling items of .82 compares favorably with the value of .81 obtained by the standardization group.

Sentence Sensitivity

Items evaluating sentence sensitivity operate by presenting to students with four similar sentences which differ primarily in terms of punctuation and syntax. The student is required to choose the sentence which "expresses the idea best." USVI examinees showed strong achievement along this

objective with a mean p value of .82 on these items compared to .81 obtained by the standardization group.

Paragraph Arrangement

This section of the English test evaluates the examinee's ability to organize a logical paragraph by ordering four sentences which are presented into their proper sequence. USVI students showed adequate ability to accomplish this task with a mean p value of .73 compared to a mean p value of .77 on these items when attempted by the standardization group.

Mathematics

All items appear to discriminate adequately with the following exceptions.

Items 1, 2, 5, and 30 - These were all very easy items for the students in the USVI sample (p=.92, .94, .92, and .90, respectively). As a result of this, all students regardless of general mathematics achievement level, were able to score correctly on these items.

Item 11 - This was a moderately difficult item (p=.63) which asked students to choose the name which "best describes a closed figure with four congruent sides and four congruent angles." The choices were, "rectangle", "rhombus", "quadrilateral", and "square." It was noted that examinees, particularly in the high achieving group, who missed this item tended to choose "quadrilateral" rather than "square" as the correct answer. The described figure is, indeed,

a quadrilateral. It is also a rectangle and a rhombus. It is obviously the intent of the author of the item that examinees realize that the four sides are equal (congruent), the angles are equal and must each measure 90 degrees, and therefore, the opposite sides must be parallel. The requirement that the student choose the "best" description may be ambiguous. It is possible that some students may not know what "best" is in this case. The item author obviously means, "most precise", but "best" could also mean "most generalizable", which would make "quadrilateral" the correct answer. It is not certain whether a student choosing "quadrilateral" did so because he or she was unable to go through the thought process described above or because he or she defined "best" in a different manner than the author. The item may be ambiguous because of this and it is uncertain what information concerning student achievement it provides.

Summary of Student Skills

Numbers, Symbols, and Sets

Students in the St. Thomas/St. John school district showed adequate achievement levels in the following objectives:

a) use of expanded notation

b) converting standard numerals to word form

c) determining place value

d) defining "prime number"

They showed weaknesses in:

a) determining relationships between two sets

b) determining the value of a linear expression with two unknowns when given the values of the unknowns

c) interpreting Venn diagrams.

St. Croix school district examinees demonstrated adequate achievement in all the objectives listed above with the exception of identifying the relationship between two sets.

Number Properties and Operations - Whole Numbers

St. Thomas/St. John school district students showed adequate achievement levels on the following objectives:

a) multiplication of common fractions

b) reducing fractions to their lowest terms

c) locating decimal points in products

d) subtracting decimal fractions

These examinees showed weaknesses in achievement on the following objectives:

a) finding decimal equivalents of fractions

b) adding common fractions with reducing

c) subtracting common fractions

d) determining relative values of fractions

e) adding mixed numbers

St. Croix school district students demonstrated adequate achievement on the following objectives:

a) multiplication of common fractions

b) reducing fractions to lowest terms

c) locating decimal points in products

d) subtracting decimal fractions

They showed weaknesses in achievement in the following areas:

a) finding decimal equivalents of fractions

b) adding and subtracting fractions

c) determining relative values of fractions

d) adding mixed numbers

Number Properties and Operations - Integers and Exponents

St. Thomas/St. John students exhibited adequate achievement on all objectives in this section of the test except determining the distance between two points on a number line.

St. Croix school district examinees showed adequate achievement on all objectives on this section of the test except the ability to add negative integers.

Mathematical Sentences

USVI students showed adequate achievement on all objectives evaluated in this section of the test.

Geometry and Measurement

St. Thomas/St. John school district examinees showed weaknesses in determining the meaning of metric prefixes while St. Croix district examinees achieved adequately on both objectives in this section of the test.

Ratio and Percent

USVI students tested achieved adequately on all objectives dealing with ratio and percent except on the objective of finding a total when a part and its percent of the total is given.

Graphs, Probability, and Statistics

Students in both USVI school districts showed adequate achievement on the following objectives:

a) identifying percents represented on a circle graph

b) finding averages of groups of numbers.

These students showed weaknesses on the following objectives:

a) computing percents represented on circle graphs

b) computing the amounts represented by percents on circle graphs.

Mathematical Reasoning

St. Thomas/St. John district examinees showed adequate achievement levels on the following objectives:

a) solving three step problems involving money

b) determing missing data needed to solve a problem.

They showed weaker achievement on these objectives:

a) stating the solution of a word problem as a mathematical expression

b) solving word problems by deduction

St. Croix district students achieved at adequate levels on all of the objectives listed above with the exception of the

ability to state solutions to word problems as mathematical expressions.

Grade 12 - TASK II Level

The Stanford Achievement Test - Test of Academic Skills, Level II consists of three tests with the number of items listed below:

1. Reading - 78 items
2. English - 69 items
3. Mathematics - 48 items

The means, standard deviations, and reliability estimates for the raw scores on these tests, as previously reported (see Bliss, 1982) are presented below.

Test	Mean Raw Score	Standard Deviation	KR-20 Reliability	Standard Error
		Total USVI Sample		
Reading	43.9	13.8	.93	3.65
English	25.3	8.3	.91	3.01
Mathematics	46.8	11.2	.87	3.36
		St. Thomas/St. John Sample		
Reading	40.6	12.3	.91	3.69
English	24.6	7.4	.90	2.98
Mathematics	45.6	10.9	.84	3.45
		St. Croix Sample		
Reading	48.2	15.2	.95	3.40
English	26.3	9.4	.91	2.97
Mathematics	48.4	11.5	.90	3.44

Following are the item analyses and summaries of student achievement on the objectives evaluated by each of the tests.

Reading

Item Analysis

As on the tenth and eighth grade levels, it was noted that a high proportion of students omitted later items on the reading test. Table 4 presents the proportion of examinees omitting each item. It should be noted that items 52 through 78 make up Part B of the test and that these items are timed separately. This accounts for the drop in omits beginning with item 52. Again, this phenomenon very likely indicates that large numbers of U.S. Virgin Islands students did not have time to attempt later items on the test and the validity of the results on these later items is definitely questionable. Because of this, only items which less than 33% of the examinees omitted will be considered in the item analysis and summary of student skills.

All items which 67% or more of the examinees did not omit discriminated at a satisfactory level except the following items.

Item # 1, 3, 9, 11, and 15 - These were particularly easy items for the USVI sample (p=.95, .84, .94, .83, and .97, respectively) and low achieving students were just as likely to score correctly on these as were high achievers. This accounts for the low discrimination indices.

Table 4

Proportions of the Twelfth Grade Sample
Omitting Reading Test Items
(Percents)

ITEM #	USVI	STT/STJ	STX
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	2	0	4
6	1	1	0
7	1	1	0
8	1	0	2
9	0	0	0
10	0	0	0
11	1	1	0
12	0	0	0
13	0	0	0
14	0	0	0
15	0	0	0
16	2	1	2
17	2	3	0
18	2	3	2
19	2	1	2
20	2	1	4
21	2	1	2
22	4	3	5

Table 4 Continued

23	6	5	7
24	6	8	4
25	7	5	9
26	6	5	7
27	9	9	9
28	10	9	11
29	14	15	13
30	14	15	13
31	14	15	13
32	16	16	16
33	16	15	16
34	18	16	20
35	18	16	20
36	23	23	24
37	29	27	33
38	29	24	35
39	30	26	36
40	35	34	36
41	33	28	38
42	34	30	40
43	28	30	25
44	31	31	31
45	32	31	33
45	32	31	33
46	34	34	35
47	37	35	40

Table 4 Continued

48	37	35	40
49	38	34	44
50	38	34	44
51	39	35	44
52	5	8	2
53	10	14	5
54	6	9	2
55	9	11	5
56	8	11	4
57	7	9	4
58	6	9	2
59	7	11	2
60	9	12	4
61	11	16	4
62	10	16	2
63	12	19	4
64	10	15	4
65	13	19	5
66	17	24	7
67	15	23	4
68	16	24	5
69	16	24	4
70	16	24	7
71	19	24	13
72	15	23	4
73	16	24	4

Table 4 Continued

74	16	24	4
75	19	26	11
76	17	26	5
77	19	26	11
78	18	28	4

Item # 18 - This moderately difficult item (p=.63) showed particularly low discrimination in the St. Thomas/St. John sample (d= .05 with p=.57). The item involved "subject" as it was used in reading passage about music. The response "object" was chosen by most of the high achieving students who missed this item rather than "theme," which was the correct response. Low achievers who scored incorrectly in the St. Thomas/St. John sample answered more or less randomly. High achieving students who were not certain of the correct answer may have been distracted by the alternative "object" since it is used often in language in connection with the term "subject." Low achieving students were likely not to have perceived this relationship and to have answered more randomly with some students choosing the correct response by chance. This could account for the low discrimination observed in this item.

Item # 7 - This was a very difficult item for St. Thomas/St. John students (p=.07), but was somewhat easier for St. Croix school district students (p=.22). The poorer discrimination in the St. Thomas/St. John district resulted from the fact that fewer high achieving students in that district scored correctly on this item. The proportions of low achieving students scoring correctly was approximately the same in both distracts. The item asked examinees to choose an example of a "geographic building

material", defined in the reading as a material which "uses local resources . . . and suits the local climate." Choices were:

a) stone for a pyramid in Egypt

b) snow for an igloo in Northern Alaska

c) adobe for a museum in New York City

d) a steel girder for an office building in Pittsburgh

The response considered correct was "b." Most St. Thomas/St. John district students chose "c" as the correct response while those St. Croix examinees in the high achieving group who scored incorrectly chose more or less evenly between "a" and "c." Adobe is the only material specifically identified as a geographic building material in the passage (although only in reference to to a community in New Mexico and the "hot, dry southwestern part of the United States." Students may have seen adobe in the passage and chosen response "c" on the basis of this, disregarding the location of the museum. On the other hand, they may not have been aware of New York City geography and climate. Neither explanation takes into account the differences observed in the behavior of the St. Croix sample. In general, this may be a poor item for a number of reasons. It requires examinees to have a strong knowledge of geographic facts not given in the readings. This may make it a good geography test item, but it is uncertain whether a student answering incorrectly did so due to lack of geographical knowledge or lack of reading skills.

In addition, the item may really have two correct answers. Stone for a pyramid might well be considered a geographical building material since it is natural to Egypt and suits the local climate there. Hence, the results on this item should be examined with caution.

Summary of Student Skills

Global Meaning

Students in the USVI sample showed some weaknesses in the area of global meaning determination. However, they seemed to score well on earlier, easier items. Examinees seem to score better, in absolute terms and in relation to the standardization group, on earlier items dealing with global meaning. On the four items with omit rates of less than 33%, the mean p value was .56 compared to .65 for the standardization group.

Meaning of Explicit Detail

USVI students appear to have achieved quite well on objectives dealing with determining explicit meaning from readings. The mean p value for the ten items measuring these skills was .74 compared to .75 for the standardization group. On three items the proportion of USVI students scoring correctly was significantly greater than the proportion of the standardization group.

Meaning of Implicit Detail

The students in the USVI sample seem to have achieved adequately in objectives dealing with the determination of facts which are implicit in readings. On items which an

appropriate number of students were able to attempt to respond to the mean p values were .67 in the St. Thomas/St. John district and .74 in the St. Croix district compared to .74 in the standardization group. Again, it was observed that local students tend to do less well on items ocurring later in the test.

Inference and Logical Analysis

USVI students showed some weaknesses in the area of inference and analysis. Of the items which were appropriate to examine, the mean p values were .51 for the St. Thomas/ St. John school district and .56 for the St. Croix school district compared to .64 for the standardization group. While the St. Croix group's achievement level could be said to be adequate, there should be some concern about the level of achievement in the St. Thomas/St. John district.

Word Meaning

Word mean involves vocabulary skills. Here, marked differences were noted between examinees in the two school districts. St. Croix district students showed considerably higher achievement (mean p value = .68) than the St. Thomas/ St. John district examinees (p=.52). It was noted that the St. Thomas/St. John examinees omitted items considerably more often than St. Croix examinees. For example, 4% of the St. Croix examinees omitted the last item (#78) on the test compared to 28% in the St. Thomas/St. John district. This difference seems to be due to one St. Thomas/St. John class

where 35% of the examinees omitted the item compared to 14% in the other. Since scores on the other parts of the battery are not greatly different between these two classes, it is possible that there was a problem in test administration in the former class. Such things as inadequate instruction on this part of the test or errors in timing could be the cause of this phenomenon. At any rate, the St. Thomas/St. John scores validity are open to question.

The St. Croix examinees show good achievement on these items with their mean p value (.68) comparing favorably with the standardization mean p value of .70.

English

Item Analysis

All items discriminated adequately except the following:

Items # 3, 31, 37, 42, 46, and 52 - These were particularly easy items for the USVI sample of examinees (p= .96. .93, .96, .96, .96, and .92, respecitively). This should account for the low discrimination indices observed on these items since low achieving students were almost as likely to score correctly on these items as were high achieving students because they were easy.

Item # 18 - This item discriminated adequately for the St. Thomas/St. John school district examinees, but poorly for the St.Croix examinees. This was due to the fact that St. Croix

examinees found the item much easier (p=.95) than the St. Thomas/St. John examinees (p=.86) and this produced the low level of discrimination typically found on items with high difficulty indices.

Item 54 - This was a moderately easy item for USVI examinees (p=.83). It was difficult to see why this item discriminated as poorly as it did. There appeared to be no pattern of incorrect responses and no extraneous clues in the item itself. The item appeared to be clear and worded simply.

Summary of Student Skills

Learning Skills

On objectives concerned with dictionary usage USVI students tested appeared to have difficulty using the pronounciation key to identify words containing given vowel sounds. St. Croix examinees also showed some weakness in their ability to locate entry words in a dictionary. On all other dictionary skills tested, the USVI sample showed adequate achievement.

Students in the St. Croix sample showed exceptionally strong achievement on objectives dealing with the use of reference sources. In fact, students in this sample had proportions answering correctly significantly higher than the standardization group on all items testing these objectives.

St. Thomas/St. John examinees seem to have achieved adequately on all objectives dealing with the use of reference sources except the ability to identify the use of a thesaurus.

On objectives dealing with knowledge of the use and structure of English USVI examinees showed strong achievement on the following:

a) identification of a prepositional phrase

b) identification of an adverb forming morpheme

They demonstrated weaknesses on the following objectives:

a) identifying an appropriate place for slang expressions

b) identifying an appropriate closing for a formal letter

c) identifying a standard English sentence

d) identifying a sentence containing a prepositional phrase.

Usage Conventions - Capitalization

Examinees in both districts showed weaknesses in their knowledge of capitalization conventions for common nouns. In general, they tended to capitalize common nouns that should have been written using lower case initial letters.

St. Thomas/St. John district students showed weaknesses in the knowledge that nationalities should be capitalized while St. Croix district students appeared to have achieved adequately along this objective.

Examinees in both districts appeared to have little difficulty recognizing that days of the week should be capitalized.

Usage Conventions - Punctuation

Examinees in the USVI sample showed adequate achievement

on all items testing the appropriate use of punctuation.

Usage Conventions - Parts of Speech

St. Thomas/St. John school district examinees demonstrated adequate achievement on all objectives dealing with usage of parts of speech tested with the exception of:

a) verb redundancy

b) the use of a participle

St. Croix school district students showed adequate achievement on all these objectives except the use of a participle.

Spelling

USVI examinees showed adequate achievement in spelling with a mean p value of .70 compared to .72 for the standardization group. On the 15 items testing spelling skills they had significantly higher proportions scored correctly than the standardization group on three items and significantly lower proportions scored correctly on only three items.

Sentence Sensitivity

Students in the USVI sample achieved well on items designed to test these objectives. Their proportions scored correctly were as high or higher than the standardization group on all items with a mean p of .77 compared to .76 for the standardization group.

Paragraph Arrangement

USVI examinees achieved adequately on objectives designed to test their abilities to arrange paragraphs with a mean p value of .61 compared to .67 for the standardization group.

Mathematics

Item Analysis

All items discriminated adequately with the following exceptions:

Item # 1 and 2 - These were particularly easy items for the USVI sample (p=.94 and .88 respectively). As a result of this phenomenon it was apparent that low achieving students had almost the same probability of scoring correctly on these items as did high achieving students. This would account for the low discrimination power which was observed.

Item # 44 - This was a particularly difficult item for students in the USVI sample (p=.12). This item is notable because of the moderate negative discrimination in the St. Croix group. However, it was noted that 45% of the St. Thomas/ St. John examinees and 27% of the St. Croix examinees omitted this item (it was near the end of the 48 item test) and it is believed that any attempt to interpret these results may be reading more validity into the item than is warranted. As on the reading test, there may be a timing difficulty involved with this test, although it appears less apparent with this test. For instance, the rise in the proportion of omits is not nearly as smooth as the rise on the reading test (See Table 5). It seems as if students have had a chance to attempt most items, but left those they had difficulty answering blank. This

Table 5

Proportions of the Twelfth Grade Sample
Omitting Mathematics Test Items
(Percents)

ITEM #	USVI	STT/STJ	STX
1	0	0	0
2	0	0	0
3	0	0	0
4	7	5	9
5	0	0	0
6	0	0	0
7	4	0	9
8	0	0	0
9	5	5	4
10	8	8	7
11	0	0	0
12	0	0	0
13	0	0	0
14	1	1	0
15	2	1	4
16	1	0	2
17	2	0	4
18	2	0	5
19	2	1	2
20	3	0	7
21	0	0	0
22	2	3	2

Table 5 Continued

23	0	0	0
24	2	1	2
25	1	0	2
26	5	5	5
27	2	4	0
28	4	5	2
29	6	7	5
30	4	5	2
31	9	15	0
32	5	5	5
33	5	5	5
34	4	5	2
35	4	5	2
36	4	5	2
37	8	9	5
38	11	12	9
39	10	15	4
40	23	28	16
41	19	24	13
42	18	27	5
43	20	31	5
44	37	45	27
45	32	41	20
46	29	43	11
47	26	39	7
48	33	46	15

seems apparent, for instance, in the St. Croix sample where 27% of the examinees omitted item number 44, but only 7% omitted item # 47. Again, a larger number of omits was observed for the St. Thomas/St. John sample than for the St. Croix sample. Since the mean number correct for these two samples differed by only 1.7 points (See Bliss, 1982) it is unlikely that the differences in omits were due to differences in math ability between groups. Instead, as in the reading test, it is more likely the differences observed were due to administration problems in the St. Thomas/St. John sample.

Summary of Student Skills

Numbers, Symbols and Sets

Examinees in the St. Thomas/St. John school district showed adequate achievement on the following objectives:

a) identification of place value of a digit in a decimal

b) identification of the relationship between two sets.

They showed weaknesses on the following objectives:

a) rounding decimals

b) interpreting Venn diagrams

c) interpreting absolute values of a number.

St. Croix district examinees showed precisely the opposite strengths and weaknesses. They appeared to show adequate achievement on the three objectives on which the St. Thomas/St. John examinees were weak and to show weak achievement on

the two objectives on which the former group of examinees displayed adequate achievement.

Number , Properties and Operations - Whole Numbers

St. Thomas/St. John district examinees displayed adequate achievement on all objectives in this grouping.

St. Croix district students showed adequate achievement on the following objectives:

a) adding a broken column of three numbers

b) subtracting with renaming

c) multiplying involving zeroes

d) multiplication of prime numbers

e) application of the multiplicative property over addition.

They showed weakness in the area of identifying common multiples of prime numbers.

Number Properties and Operations - Common and Decimal Fractions

St. Thomas/St. John district examinees showed adequate achievement on all objectives in this area except:

a) adding fractions with reduction

b) estimating the quotient of decimal division

c) identifying the least common denominator of literal fractions

St. Croix district examinees showed adequate achievement on all objectives in this area except:

a) adding fractions with reduction

b) estimating the quotient in decimal division

c) subtracting common fractions with unlike denominators after renaming.

Number Properties and Operations - Integers and Exponents

USVI examinees showed adequate achievement on all objectives tested in this area.

Mathematical Sentences

St. Thomas/St. John school district examinees showed adequate achievement on all objectives tested in this area except the ability to translate verbal problems into mathematical equations.

St. Croix district students showed adequate achievement on all objectives tested in this item grouping.

Geometry and Measurement

There were substantial differences in achievement in this area between examinees in the two school districts.

St. Thomas/St. John district students displayed adequate achievement on the ability to approximate the English equivalent of a linear metric unit. They showed weak achievement on the following objectives:

a) identifying characteristics of common polygons

b) determining the relative areas of polygons, given their dimensions

c) naming the angles that result from bisecting a right angle.

d) solving one step word problems that involve converting between English linear units.

St. Croix district examinees showed particularly strong achievement on the following objectives:

a) identifying characteristics of common polygons

b) naming the angles that result from bisecting a right angle

They showed adequate achievement on these objectives:

a) determining the relative areas of polygons, given their dimensions

b) approximating the English equivalent of a linear metric unit

These students showed weakness in solving one step word problems that involve converting between English linear units.

Ratio and Percent

St. Thomas/St. John district examinees displayed adequate achievement on all objectives tested in this area.

St. Croix district students showed adequate achievement on the objectives tested in this item grouping except the ability to solve rate problems.

Graphs, Probability, and Statistics

St. Thomas/St. John district examinees demonstrated adequate achievement in the ability to use tables to make comparisons. They showed weaknesses in the following areas:

a) solving word problems by finding the average of two numbers

b) finding the average of positive and negative numbers

c) solving word problems by interpreting a table to find percent success

d) identifying the probability of a particular event occurring.

Two other items had omits above 40% and no attempt will be made to derive information from them.

St. Croix district examinees showed adequate achievement on the following objectives:

a) finds the average of a positive and negative number
b) using tables to make comparisons
c) choosing a set of ordered pairs that would be on the same line as two given points.

They showed weaknesses on the following objectives:

a) solving word problems by finding the average of two numbers
b) solving word problems by interpreting a table to find percent of success
c) identifying the probability of a particular event occurring
d) identifying coordinates on a graph of a linear equation
e) choosing a set of ordered pairs that would be on the same line as two given points.

Mathematical Reasoning

USVI examinees displayed adequate achievement on all objectives tested in this item group. In the St. Thomas/ St. John sample, one item had 46% omitted responses (item # 48) and was not used in this analysis.

Discussion

A detailed summary of student basic skills across grades and recommendations based on the results of the achievement testing carried out will, of necessity, have to be put off until the results of elementary school testing are analyzed and reported. Nevertheless, there are some facts brought out by the analysis of the secondary school examinee scores which are worth, at least tentatively, touching upon.

Most striking is the time situation on the reading comprehension tests. On all three levels tested it was obvious that many students did not have sufficient time to attempt all the items on these tests. Teachers and counselors were instructed to strictly adhere to the time limits specified by the test publishers. These times had been empirically determined to be sufficient for all students to complete the test during the standardization studies described in Technical Report # 1. It appears that these time limits may not be appropriate for USVI students who may be slower, more deliberate readers. This characteristic of local students is not necessarily undesirable provided students end up understanding what they read. Nevertheless, it may account for low reading scores since most standardized tests count omits as incorrect answers. From a technical point of view, the result of this phenomenon may be that students are taking a timed test rather than a power test and this brings up numerous problems in score interpretation.

It is notable that time seems to be sufficient on most other tests. No attempt is made here to describe the causes of this phenomenon. Many alternative hypothesis present themselves. However, it is important that research be conducted to establish the nature and causes of this difficulty.

Second, there appears to be a pattern that can, at least tentatively, be observed in student achievement. Using the taxonomy of educational objectives in the cognitive domain (Bloom, 1956) as a theoretical framework, some sense can be made out of this pattern. U.S. Virgin Islands students generally appear to achieve at least adequately on objectives that would be classified in the lower three categories of the taxonomy (i.e. knowledge, comprehension, and application). They achieve less well on objectives that would be classified in the upper three categories (analysis, synthesis, and evaluation). Again the reading comprehension tests exhibit this phenomenon quite clearly. USVI examinees achieve adequately on items which test their abilities to understand facts that are explicitly stated in the text (comprehension level objectives), but are weak in determining implicit, global, and contextual meanings (analysis and synthesis objectives). These students seem to achieve well on objectives which can be achieved by memorization (knowledge objectives) or the application of a general procedure, rule or algorithm (comprehension and application level). They scored extremely well in spelling and single step mathematical computations. They scored lower on items requiring

them to organize data, choose information relevant to solving problems, and to choose a series of rules and/or algorithms to use in solving problems (analysis, synthesis, and evaluation objectives). Note the weaker achievement in mathematics applications problems, multistep mathematical computations, and the use of dictionary entries, for example. Again, there is no attempt here to determine the causes of this phenomenon, but if analysis of the elementary achievement tests substantiates this tentative observation, further research into this phenomenon is important.

Third, U.S. Virgin Islands students in the grade levels tested displayed weakness in their knowledge of standard English vocabulary.

Finally, there is the issue of test taking skills. One particular phenomenon was very apparent. Many examinees omitted items which they had a chance to attempt to answer. Although it was not explicitly stated in the directions, the score students received was the number of items they answered correctly. There was no correction for guessing and an omit was counted the same as a wrong answer. This should have been explained in the directions and the lack of such an explanation shows a deficit in the test publishers' thoroughness in test construction. Nevertheless, most tests students take use this "number right" scoring procedure (exceptions are the Preliminary Scholastic Aptitude Test and the Scholastic Aptitude Test which are not given earlier than grade eleven) and the best strategy

for examinees wishing to maximize their scores when confronted with an item they did not know the answer to would have been to try to discard alternatives which they know were wrong and to guess randomly from the remaining alternatives or, if they could identify no incorrect answers, to simply take a random guess. Many examinees seemed to be unaware of this strategy or unwilling or unable to use it.

References

Allen, M.J. & Yen, W.M. Introduction to measurement theory. Monterey, CA: Brooks/Cole Publishing Co., 1979.

Bliss, L.B. Virgin Islands of the United States public school basic skills achievement survey, Technical Report #1: Validation of the use of the Stanford Achievement Test with U.S.V.I. students. St. Thomas, USVI: Caribbean Research Institute, College of the Virgin Islands, 1982.

Bloom, B. (Ed.) Taxonomy of educational objectives, Handbook I: Cognitive domain. New York: Longmans, Green & Co., 1956.

Cooper, V. Personal communication, June 7, 1982.

Cureton, E.E. The upper and lower twenty-seven per cent rule. Psychometrika, 1957, 22, 293-296.

Kelly, T.L. The selection of upper and lower groups for the validation of test items. Journal of Educational Psychology, 1939, 30, 17-24.